



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/848,982	05/03/2001	Ted E. Dunning	22227-05479	8782
758	7590	06/17/2004	EXAMINER	
FENWICK & WEST LLP SILICON VALLEY CENTER 801 CALIFORNIA STREET MOUNTAIN VIEW, CA 94041			WONG, LESLIE	
		ART UNIT	PAPER NUMBER	
		2177	20	
DATE MAILED: 06/17/2004				

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)
	09/848,982	DUNNING ET AL. <i>hr</i>
	Examiner	Art Unit
	Leslie Wong	2177

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extension of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 29 March 2004.
 2a) This action is FINAL. 2b) This action is non-final.
 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1-35 is/are pending in the application.
 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
 5) Claim(s) _____ is/are allowed.
 6) Claim(s) 1-35 is/are rejected.
 7) Claim(s) _____ is/are objected to.
 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.
 10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) Notice of References Cited (PTO-892)
 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)
 3) Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
 Paper No(s)/Mail Date _____

4) Interview Summary (PTO-413)
 Paper No(s)/Mail Date. _____
 5) Notice of Informal Patent Application (PTO-152)
 6) Other: _____

DETAILED ACTION

Priority

1. The Applicant's claim to domestic priority under 35 U.S.C. §119(e), as a provisional of application serial number **60/201622**, filed on **03 May 2000**, is acknowledged.

Remarks

2. The "Currently Amended" statuses for claims 12, 23, and 34 appear to be typographical errors as there have been no modifications made to the immediate prior version of the claims.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to

consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

4. Claims 1-35 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Damashek** (U.S. Patent 5,418,951) in view of **Ortega et al.** (U.S. Patent Application 20020152204A1).

Regarding claims 1, 12, 23, and 34, **Damashek** teaches a computer-implemented method, system, and computer-readable medium for performing text equivalencing from a string of characters comprising:

a). **'modifying the string of characters using a predetermined set of heuristics'** as reducing multiple spaces to a single space within a string of characters and the strings of characters may also be eliminated or replaced by a user-defined character or strings of characters (col. 4, line 64 – col. 5, line 5; col. 8, line 64 col. 9, line 2);

b). **'comparing the modified string with a known string of characters in order to locate a match'** as comparing the scores for the n-grams strings between the unidentified document and the reference documents to determine the degree of similarity between the strings of the two documents (col. 5, lines 54-67; col. 4, lines 10-60);

c). **'responsive to not finding an exact match, forming a plurality of sub-strings of characters from the string of characters'** as parsing text which is written in an unidentified language into n-grams. N-grams (i.e. sub-strings) are consecutive runs of n characters where n is any positive integer greater than zero (col. 4, lines 49-56; col. 5, lines 24-30; col. 3, lines 21-24; col. 4, lines 24-27); and

d). **'using an information retrieval technique on the sub-strings of characters to determine a known string of characters equivalent to the string of characters'** as enumerating the n-grams contained in the unidentified document and comparing the result of that operation with the enumerated n-grams found in a reference document (col. 3, lines 22-34 and col. 4, lines 10-60).

b). **Damashek** does not explicitly teach a step of performing a character-by-character comparison of the strings.

Ortega et al., however, teaches a step of '**performing a character-by-character comparison of the strings**' as comparing a non-matching term to the list of related terms one-by-one using an anagram-type function which compares two character-strings and returns a numerical similarity score (¶s 0021, 0033, 0057-0064).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to combine the teachings of the cited references because **Ortega's** teaching would have allowed **Damashek's** to facilitate processing and increasing the efficiency of a search query by invoking the spelling correction process to attempt to correct the non-matching term(s) and comparing a non-matching term of a search

criteria with data in the correlation table to identify any possible replacements (¶ 0010) as suggested by **Ortega et al.** at ¶s 0054 and 0066.

Regarding claims 2, 13, 24, and 35, **Damashek** further teaches a step wherein the information retrieval technique further comprises:

- a). weighting the sub-strings (col. 5, lines 31);
- b). scoring the known string of characters (col. 8, lines 51-56); and
- c). retrieving information associated with the known string of characters with the highest score (col. 9, lines 64-66).

Regarding claims 3, 14, and 25, **Damashek** further teaches a step comprising, responsive to the highest score being greater than a first threshold, automatically accepting the known string of characters as an exact match (col. 8, lines 51-63).

Regarding claims 4, 15, and 26, **Damashek** further teaches a step comprising, responsive to the highest score being less than a second threshold and greater than a first threshold, presenting the known string of characters to a user for manual confirmation (col. 9, lines 12-14; col. 10. 45-49).

Regarding claims 5, 16, and 27, **Damashek** further teaches a step comprising, responsive to the highest score being less than a second threshold and greater than a

third threshold, presenting the known string of characters to a user to select the equivalent string of characters (col. 9, lines 12-14; col. 10, 45-49).

Regarding claims 6, 17, and 28, **Damashek** further teaches a step wherein the sub-strings of characters are 3-grams (col. 3, lines 21-24; col. 4, lines 24-27).

Regarding claims 7, 18, and 29, **Damashek** further teaches a step wherein the string of characters is selected from the group consisting of a song title, a song artist, an album name, a book title, an author's name, a book publisher, a genetic sequence, and a computer program (col. 9, lines 35-37).

Regarding claims 8, 19, and 30, **Damashek** further teaches a step wherein the predetermined set of heuristics comprises removing whitespace from the string of characters (col. 4, line 64 – col. 5, line 5).

Regarding claims 9, 20, and 31, **Damashek** further teaches a step wherein the predetermined set of heuristics comprises removing a portion of the string of characters (col. 8, line 64 – col. 9, line 10).

Regarding claims 10, 21, and 32, **Damashek** further teaches a step wherein the predetermined set of heuristics comprises replacing a symbol in the string of characters with an alternate representation for the symbol (col. 4, line 64 – col. 5, line 5).

Regarding claims 11, 22, and 33, **Damashek** further teaches a step wherein storing an indication (i.e., similarity score) that the string of characters is the equivalent of the known string of characters (col. 8, lines 51-56).

Response to Argument

5. Applicant's arguments regarding the Haimowitz reference have been considered but are moot in view of the new ground(s) of rejection.

Applicants argue that the goal of Damashek is to determine a language or topic for the unidentified document or query and that nowhere in Damashek is there any attempt to determine a text equivalent for a document or query, or to perform any type of string matching or comparison. In response to the preceding arguments, the Examiner respectfully submits that Damashek teaches finding a match for a character string by comparing it with known text (i.e., reference documents) (col. 5, lines 10-11). The reference documents are parsed into sub-strings (n-grams) for each reference document. Weights are assigned to each unique sub-string (i.e., n-gram). The weight is determined by the relative frequency of occurrence of that n-gram in the reference document (col. 5, 24-30). The string that the system attempts to find a match for is also parsed into a list of unique n-grams and weight is also assigned to each n-gram. The string is then compared to each of the known strings by scoring the string against the known strings (i.e., reference document). The score for the string with respect to the

known strings indicates the degree of similarity between the two strings (col. 5, lines 54-60). Although Damashek teaches strings comparison using the n-grams system, in order to place the unidentified string into its proper category, that is different than Applicants' intended usage; However, the prior teaches the structure (i.e. n-grams sub-strings formations) which allows the system to find a matching or relatively similar string for the input string of characters. Damashek discloses the method for sub-strings formation using n-grams which supports the claimed limitation (i.e., "... forming a plurality of sub-strings of characters from the strings of characters..."). Hence, Damashek satisfies the claimed limitations as analyzed and discussed above.

Conclusion

6. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Martino et al. (US006009382A)

Ejiri (US005182708A)

Schulze (US006167369A)

Rosenbaum et al. (US004384329)

Register et al. (US005371807A)

Hargrave, III et al. (US006131082A)

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Leslie Wong whose telephone number is (703) 305-3018. The examiner can normally be reached on Monday to Friday 9:30am - 6:30 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John E Breene can be reached on (703) 305-9790. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).



Leslie Wong
Patent Examiner
Art Unit 2177

LW
June 14, 2004